

FIG. 1

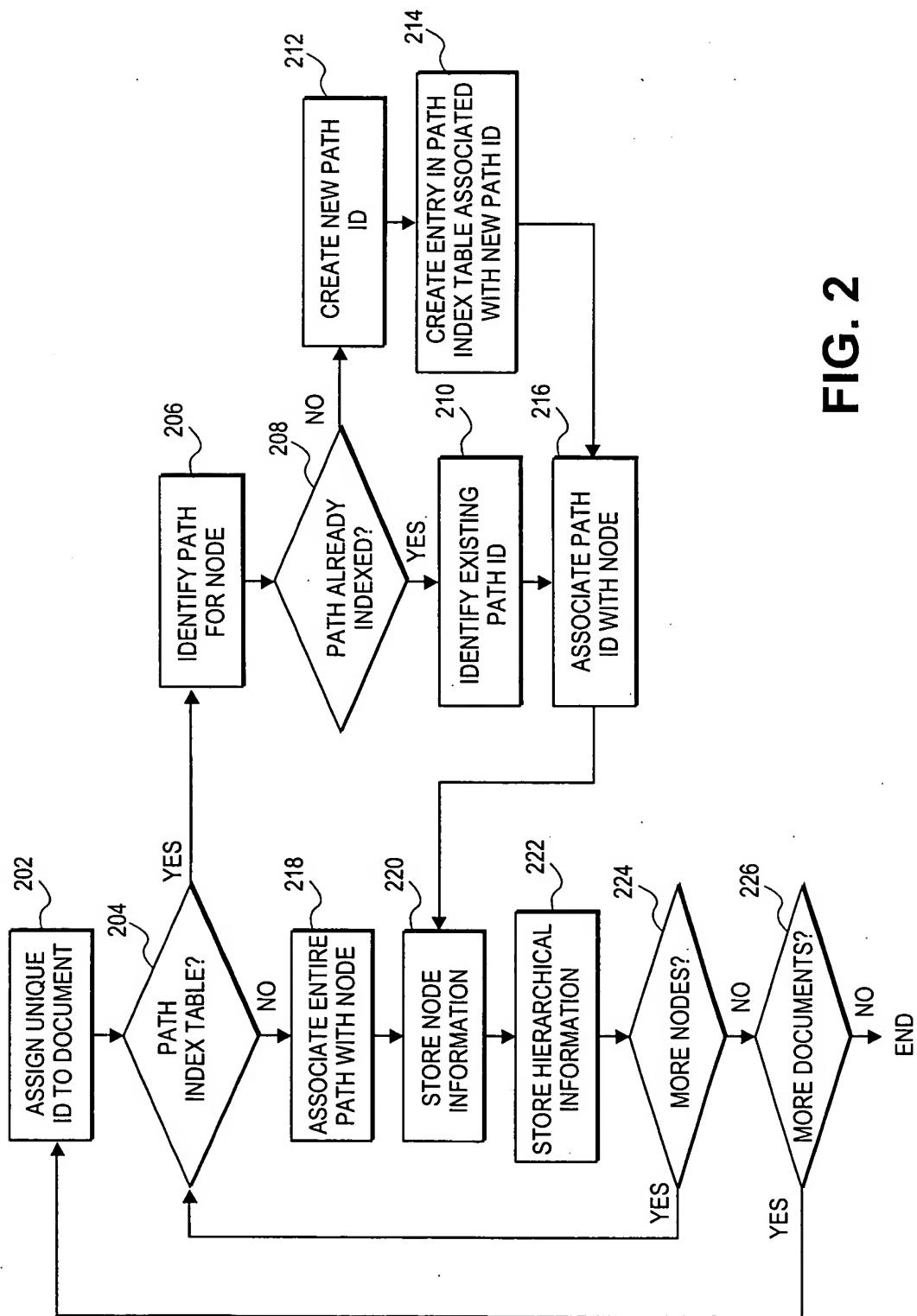
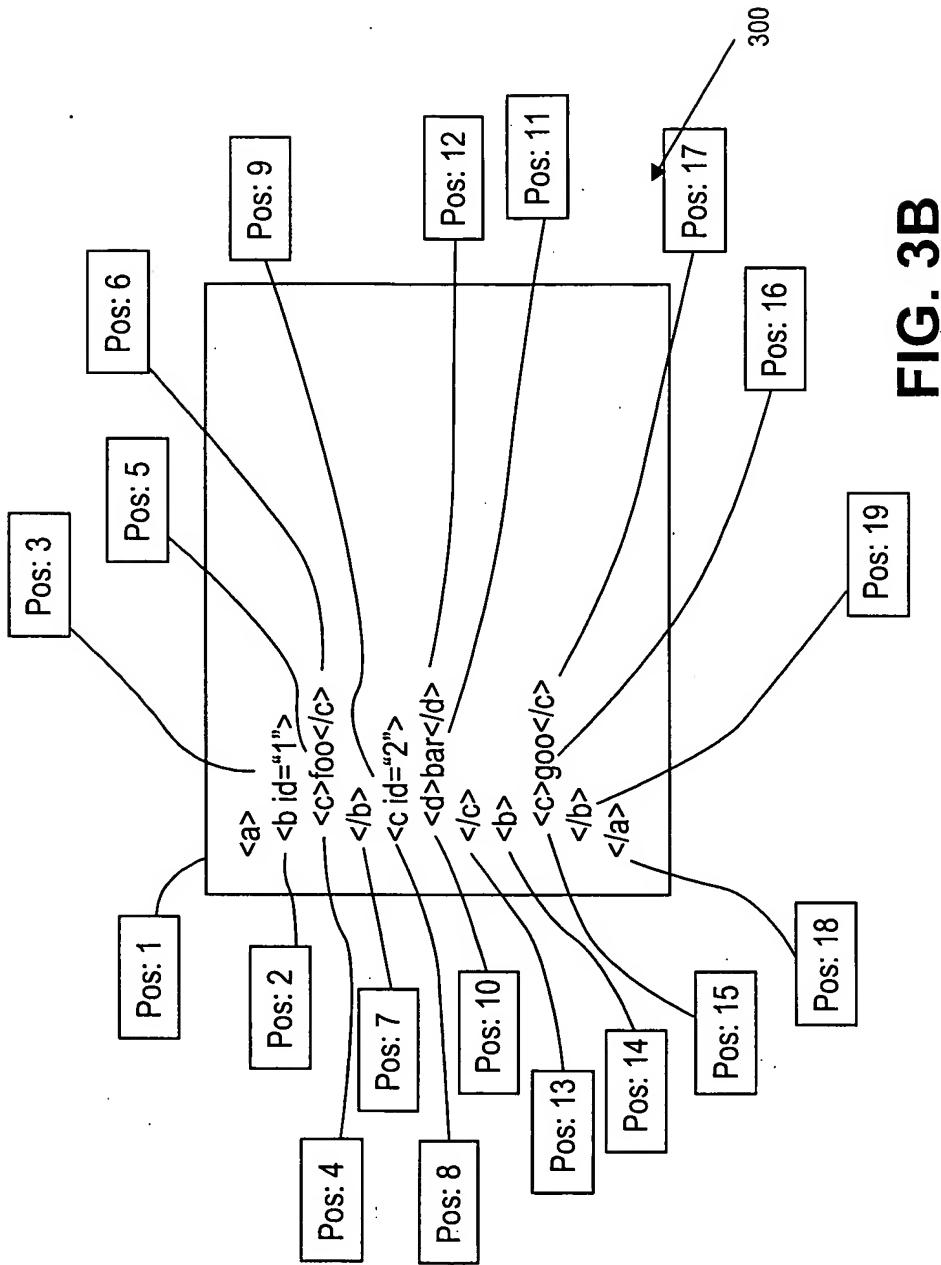


FIG. 2

```
<a>
  <b id="1">
    <c>foo</c>
  </b>
  <c id="2">
    <d>bar</d>
  </c>
  <b>
    <c>goo</c>
  </b>
</a>
```

300

FIG. 3A



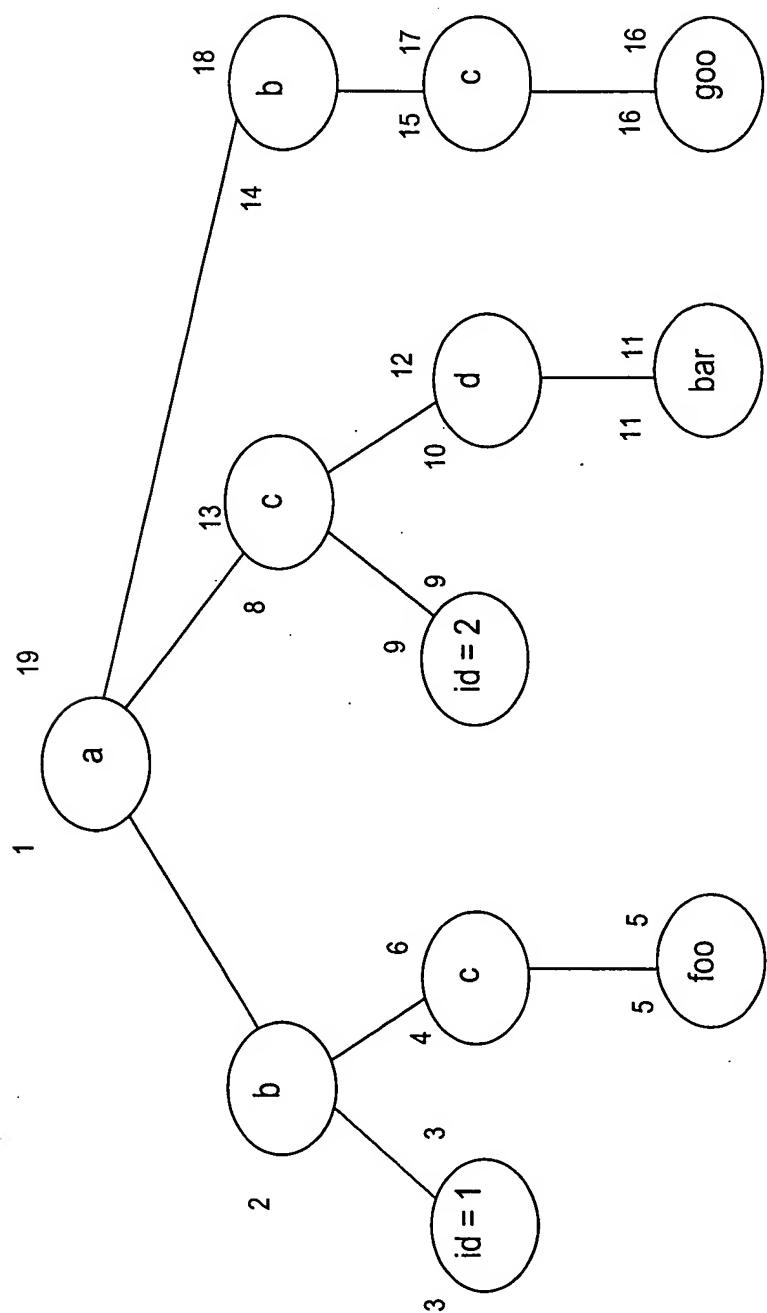


FIG. 3C

| PID | PATH | NODE NAME |
|-----|----------|-----------|
| 1 | a. | a |
| 2 | a.b. | b |
| 3 | a.b.@id. | id |
| 4 | a.b.c. | c |
| 5 | a.c. | c |
| 6 | a.c.@id. | id |
| 7 | a.c.d. | d |

402

400

FIG. 4

| DOCID | PID | STARTPOS | ENDPOS | NODELVL | NODETYPE | NODEVAL |
|-------|-----|----------|--------|---------|----------|---------|
| 1 | 1 | 1 | 19 | 1 | ELEM | |
| 1 | 2 | 2 | 7 | 2 | ELEM | |
| 1 | 3 | 3 | 3 | 3 | ATTR | 1 |
| 1 | 4 | 4 | 6 | 3 | ELEM | |
| 1 | 4 | 5 | 5 | 4 | TEXT | foo |
| 1 | 5 | 8 | 13 | 2 | ELEM | |
| 1 | 6 | 9 | 9 | 3 | ATTR | 2 |
| 1 | 7 | 10 | 12 | 3 | ELEM | |
| 1 | 7 | 11 | 11 | 4 | TEXT | bar |
| 1 | 2 | 14 | 18 | 2 | ELEM | |
| 1 | 4 | 15 | 17 | 3 | ELEM | |
| 1 | 4 | 16 | 16 | 4 | TEXT | goo |

502

500

FIG. 5

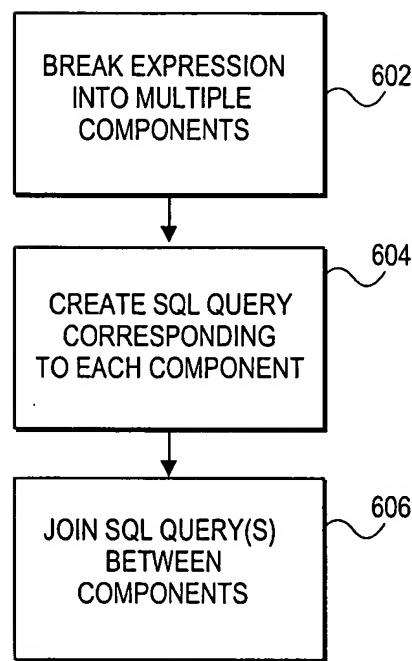


FIG. 6

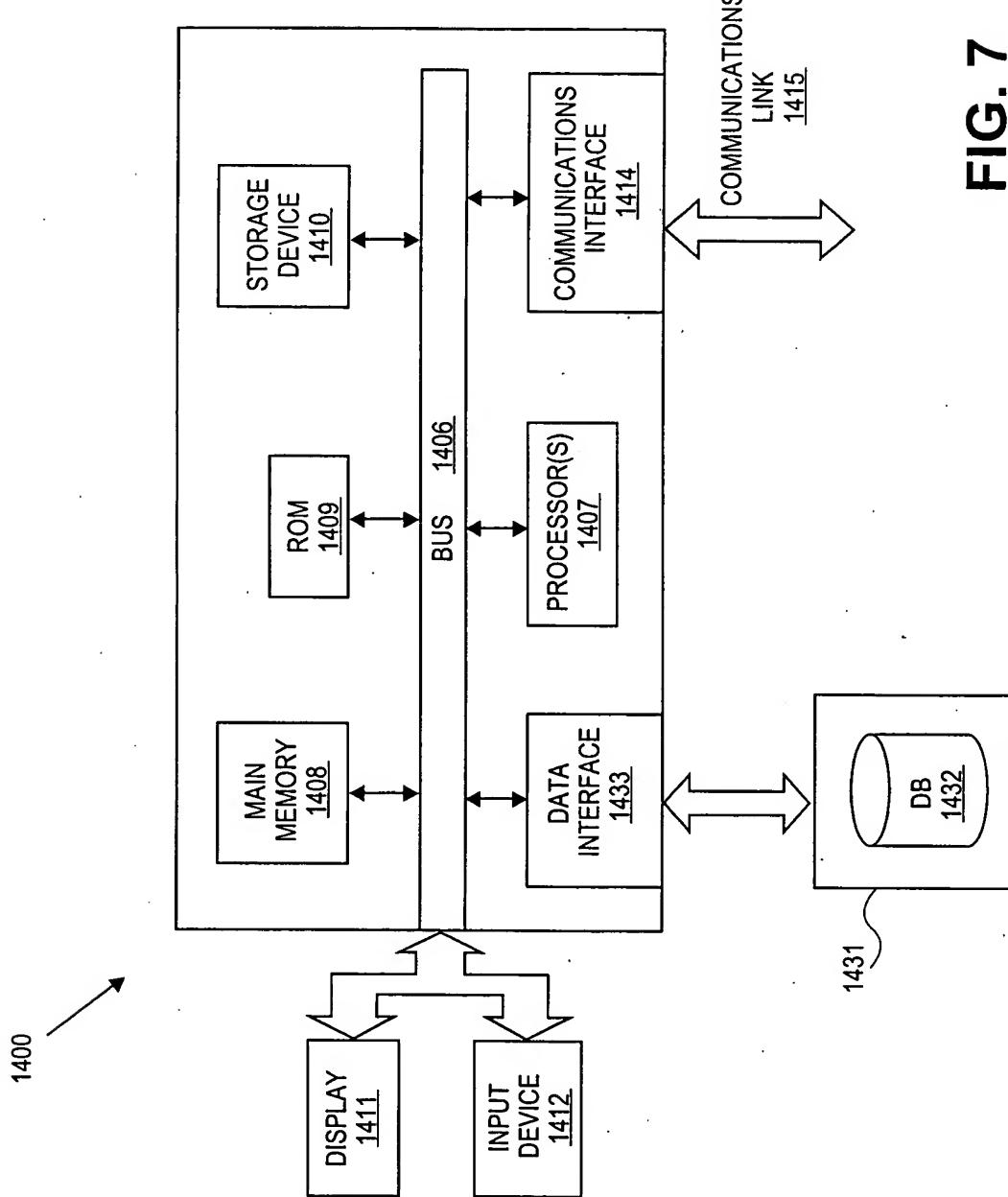


FIG. 7